



基于风格转换的无监督聚类行人重识别

张智, 毕晓君

引用本文:

张智, 毕晓君. 基于风格转换的无监督聚类行人重识别[J]. *智能系统学报*, 2021, 16(1): 48–56.

ZHANG Zhi, BI Xiaojun. Clustering approach based on style transfer for unsupervised person re-identification[J]. *CAAI Transactions on Intelligent Systems*, 2021, 16(1): 48–56.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202012014>

您可能感兴趣的其他文章

可能性匹配知识迁移原型聚类算法

Possibility-matching based knowledge transfer prototype clustering algorithm

智能系统学报. 2020, 15(5): 978–989 <https://dx.doi.org/10.11992/tis.201810028>

基于生成式对抗网络的道路交通模糊图像增强

Enhancement of blurred road-traffic images based on generative adversarial network

智能系统学报. 2020, 15(3): 491–498 <https://dx.doi.org/10.11992/tis.201903041>

生成对抗网络辅助学习的舰船目标精细识别

Fine-grained inshore ship recognition assisted by deep-learning generative adversarial networks

智能系统学报. 2020, 15(2): 296–301 <https://dx.doi.org/10.11992/tis.201901004>

基于生成对抗网络的机载遥感图像超分辨率重建

Super-resolution reconstruction of airborne remote sensing images based on the generative adversarial networks

智能系统学报. 2020, 15(1): 74–83 <https://dx.doi.org/10.11992/tis.202002002>

结合稀疏表示与约束传递的半监督谱聚类算法

A semi-supervised spectral clustering algorithm combined with sparse representation and constraint propagation

智能系统学报. 2018, 13(5): 855–862 <https://dx.doi.org/10.11992/tis.201703013>

行人重识别研究综述

Survey on pedestrian re-identification research

智能系统学报. 2017, 12(6): 770–780 <https://dx.doi.org/10.11992/tis.201706084>

微信公众平台



关注微信公众号, 获取更多资讯信息

DOI: 10.11992/tis.202012014

基于风格转换的无监督聚类行人重识别

张智¹, 毕晓君²

(1. 哈尔滨工程大学信息与通信工程学院, 黑龙江哈尔滨 150001; 2. 中央民族大学信息工程学院, 北京 100081)

摘要: 无监督行人重识别中源域与目标域间的巨大差异性是影响模型性能的最关键因素。基于聚类的无监督行人重识别方法挖掘目标域数据间的相似性, 以此缓解该问题, 但仍未消除域间差异性。本文提出一种基于风格转换的无监督聚类行人重识别方法。首先, 针对基于聚类方法的模型存在受域间差异性影响的问题, 将一种基于生成对抗网络的风格转换方法引入到聚类方法模型中, 将源域数据转换为目标域风格数据, 直接减小域间差异性, 提升模型的识别性能。其次, 针对风格转换模型的生成器存在转换尺度单一以及特征信息传递效率低的问题, 使用一种新型残差块替换原始残差块并将其引入到生成器上采样和下采样中, 形成多特征尺度转换以及信息传递效率高的生成器, 提升风格转换效果, 降低域间差异性, 进一步提升整体模型的识别效果。在 Market1501 以及 Duke-MTMC-reID 数据集上对所提的算法进行实验, 结果表明改进方法取得了更好的识别效果。

关键词: 机器视觉; 行人重识别; 无监督; 聚类; 风格转换; 生成对抗网络; 残差块; 跨域

中图分类号: TP391 **文献标志码:** A **文章编号:** 1673-4785(2021)01-0048-09

中文引用格式: 张智, 毕晓君. 基于风格转换的无监督聚类行人重识别 [J]. 智能系统学报, 2021, 16(1): 48-56.

英文引用格式: ZHANG Zhi, BI Xiaojun. Clustering approach based on style transfer for unsupervised person re-identification[J]. CAAI transactions on intelligent systems, 2021, 16(1): 48-56.

Clustering approach based on style transfer for unsupervised person re-identification

ZHANG Zhi¹, BI Xiaojun²

(1. College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China; 2. School of Information Engineering, Minzu University of China, Beijing 100081, China)

Abstract: The substantial difference between the source and target domains is the most crucial factor affecting the performance of unsupervised person re-identification models. The clustering-based unsupervised person re-identification method alleviates the problem to a certain extent by mining the similarity between the target domain, but it does not fundamentally eliminate the discrepancy between the domains. This paper proposes a clustering approach based on cross-domain style transfer for unsupervised pedestrian re-identification. First, to avoid the difference between domains in clustering-based unsupervised person re-identification models, the across-domain style transfer method based on a generative adversarial network is introduced into the clustering process. It transfers the source domain data to the target domain style data, which directly reduces the difference between domains and improves the recognition effect of the model. Second, the generator of cross-domain style transfer model has a single transfer scale and low efficiency of characteristics information transfer. A new type of residual block is proposed to replace the original residual block; then, it is inserted into the generator to achieve up-sampling and down-sampling. The specific generator has more characteristics of the scale transfer, and it transmits information more effectively. The cross-domain style transfer model can better transfer the style of the source and target domains, further reduce the difference between the two domains, and improve the recognition effect of the overall model. Extensive experiments were implemented on Market1501 and Duke-MTMC-Reid datasets to examine the proposed method, and the results showed that the proposed improved method achieved a better recognition effect.

Keywords: machine vision; pedestrian re-identification; unsupervised; clustering; style transformation; generative adversarial networks; residual block; cross domain

收稿日期: 2020-12-07.

通信作者: 毕晓君. E-mail: bixiaojun@hrbeu.edu.cn.

行人重识别是利用计算机视觉技术来判断无
视角交叠摄像头间是否存在相同行人的技术。行

人重识别技术作为提升智能监控系统自动化程度的关键技术已经成为研究热点, 该技术常与行人检测与跟踪技术结合, 广泛用于智能视频监控与公共安防的刑侦工作等。目前, 深度学习理论^[1]在各个领域都大放异彩, 现在行人重识别领域^[2]最优秀的模型都是基于深度学习的方法。

基于深度学习的行人重识别方法主要分为 2 大类, 有监督学习的行人重识别^[3-4]和无监督学习的行人重识别^[5-8]。有监督学习需要大量的标注样本指导模型训练, 由于实际应用中样本标注的成本十分昂贵, 无监督学习的行人重识别方法不需要大量的有标签样本, 因此无监督方法更具有实际应用价值。无监督方法较有监督方法更难实现, 源域和目标域间的差异性过大是目前该领域需要解决的问题。截止到目前, 对于无监督行人重识别, 国外研究刚刚起步, 而在国内还属于研究空白。因此, 无监督学习的行人重识别方法研究具有重大意义。

目前, 无监督行人重识别的研究工作中, 均使用有标签源域数据对模型进行预训练, 用无标签的目标域数据对模型进行正式训练以及测试。但是, 研究中存在域间差异性过大问题, 大幅度影响模型的识别率。2017 年, Zhong 等^[5]首次提出域连接性学习方法构建源域和目标域之间的联系, 该方法一定程度上构建了源域与目标域之间的特征空间关系, 但没有有效解决域间巨大差异性。图像风格转换 (image style transfer) 是将图像按照参考风格图像进行转换, 最终得到某种特定风格的目标图像的技术, 该技术常被应用于图像处理及视频处理, 也是无监督行人重识别领域的重要辅助技术。2018 年, Deng 等^[7]提出一种用于无监督行人重识别的风格转换模型 (similarity preserving cycleconsistent generative adversarial network, SPGAN), 该模型以循环生成对抗网络 (cycle-consistent adversarial networks, CycleGAN)^[9]为基础框架, 引入以目标域的身份作为约束的损失函数, 辅助生成器生成目标域风格的行人图片; 另外, 该模型加入孪生网络 (siamese network, SiaNet) 以及对应的一致性损失函数, 保证行人图片经过转换后其相关身份信息不丢失。但是, SPGAN 模型的生成器存在特征转换尺度单一, 且特征信息传递效率低的问题, 导致其风格转换效果不佳。2019 年, Fu 等^[8]提出一种基于聚类方法的无监督行人重识别模型 (self-similarity grouping, SSG), 通过对行人特征进行聚类操作寻找目标域数据间的相似性来对特征进行分类, 并

按照类别赋予行人特征不同的伪标签, 最后利用伪标签进行有监督学习, 取得了较好的效果。但是 SSG 模型主要通过挖掘目标域数据间的相似性来降低域间差异性过大的影响, 而没有从根本上减小域间差异性, 导致模型能力受限。

目前无监督行人重识别领域中, 没有工作通过将基于图像生成的风格转换模型引入到基于聚类方法的无监督行人重识别模型中, 来降低域间巨大差异性对聚类方法模型性能的影响。为了提升聚类方法无监督行人重识别模型的性能, 本文针对 SSG 模型存在源域和目标域之间差异性较大问题, 提出一种基于风格转换的无监督聚类行人重识别方法。首先, 将无监督行人重识别领域的风格转换模型 SPGAN 引入 SSG 模型中, 减少两域间差异性对网络识别效果的影响; 然后, 针对引入的 SPGAN 模型的生成器特征转换尺度单一以及特征信息传递效率低的缺点进行改进, 提升 SPGAN 模型的风格转换效果, 进一步减少域间差异性对 SSG 模型的影响, 提升 SSG 网络的识别效果。

1 基于风格转换的无监督聚类行人重识别

本文通过提出一种基于风格转换的无监督聚类行人重识别方法, 减少域间巨大差异性对模型性能的影响。该方法包含以下 2 个改进: 1) 将无监督行人重识别领域的风格转换模型—SPGAN 引入到基于聚类方法的无监督行人重识别模型—SSG 中, 通过风格转换方法降低域间差异性对 SSG 模型识别率的影响; 2) 为进一步减小域间巨大差异性的影响, 对 SPGAN 进行改进, 通过改进其生成器, 解决生成器中特征转换尺度单一以及特征信息传递效率低的问题, 提升 SPGAN 模型的风格转换效果, 进一步降低域间差异性对 SSG 模型识别率的影响。

本文提出的整体模型如图 1 所示, 首先, 通过基于多尺度特征转换的风格转换模型对源域图像进行风格转换, 利用转换图像替代原来的源域图像对模型进行预训练, 然后使用聚类方法对目标域数据进行聚类并分配标签, 再通过正式训练对预训练模型进行微调。

1.1 基于风格转换的 SSG 模型

无监督聚类的行人重识别模型 SSG 通过对网络提取的特征进行聚类, 并按照聚类结果为它们分配伪标签来训练模型。SSG 方法的具体步骤

如下: 1) 使用源域数据预训练 Baseline 模型; 2) 将目标域数据输入 Baseline 模型提取特征; 3) 通过聚类算法对行人特征进行聚类, 并分配伪标签; 4) 利用带有伪标签的行人特征训练模型。

域间差异性过大导致模型最终识别率的下降是目前无监督行人重识别领域的一个重大问题, SSG 模型通过聚类方法挖掘目标域数据间的相似

性降低源域和目标域间差异性过大对识别率带来的影响。由于 Baseline 模型使用源域数据进行预训练, 挖掘目标域数据间相似性仅可以侧面降低域间差异性对模型识别率的影响。为了进一步降低域间差异性过大对 SSG 模型的影响, 本文将风格转换模型 SPGAN 引入到 SSG 模型中。所提出的基于风格转换的 SSG 模型如图 2 所示。

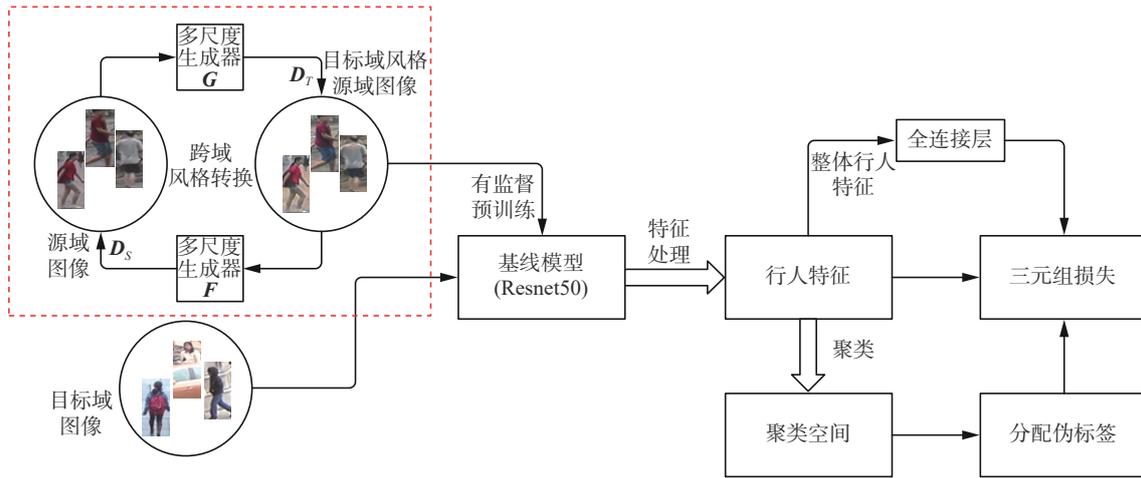


图 1 基于风格转换的聚类无监督行人重识别模型

Fig. 1 Clustering approach based on style transfer for unsupervised person re-identification model

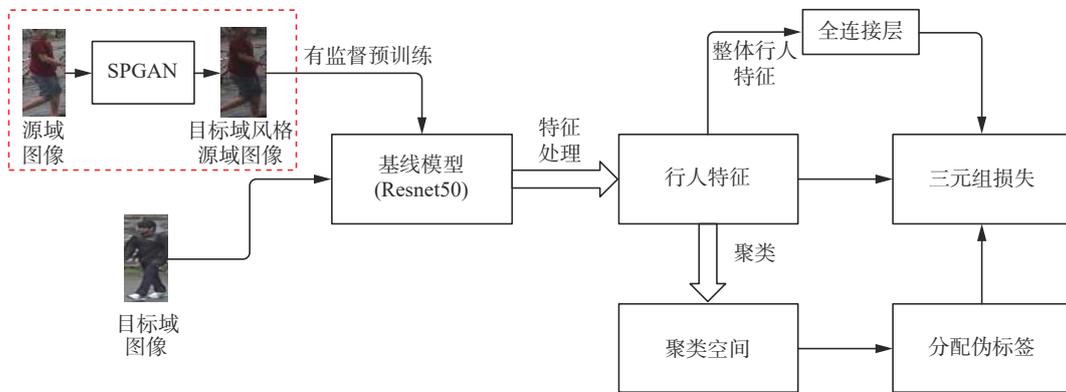


图 2 基于风格转换的 SSG 模型

Fig. 2 SSG model based on style transfer

首先, 利用 SPGAN 模型对行人重识别的源域及目标域数据集进行风格转换, 在光照、分辨率以及相机风格方面, 得到更接近目标域风格的有标签数据集; 利用转换后源域数据集预训练以 ResNet50^[10] 为基础框架的 Baseline 模型, 得到一个更加适应于目标域的 Baseline 模型。利用新的 Baseline 模型对无标签数据进行特征提取, 每个行人的特征被提取后复制为 2 份, 一份作为整体行人特征 f^i , 另一份进行横向裁剪分为上半部分行人特征 $f_{t_up}^i$ 以及下部分行人特征 $f_{t_low}^i$, 将目标域数据提取的特征重复以上步骤得到 3 组特征集

合 f^i 、 $f_{t_up}^i$ 和 $f_{t_low}^i$ 。

将 3 组特征向量分别通过基于密度的聚类方法 (density-based spatial clustering of applications with noise)^[11] 进行聚类操作, 3 组特征向量分别进行分类, 根据不同组不同类别为每个行人图像分配标签 y^i 、 $y_{t_up}^i$ 以及 $y_{t_low}^i$ 。根据 3 个特征向量的分组结果, 建立一个新的目标域数据集, 其中每张图像带有 3 个标签信息。

在新数据集上使用三元组损失函数 (batch-hard triplet loss)^[12] 损失函数对 Baseline 模型进行微调训练, 该损失函数可表示为

$$L_{\text{triplet}} = \sum_{i=1}^P \sum_{a=1}^K [a + \max_{p=1,2,\dots,K} \|x_a^{(i)} - x_p^{(i)}\|_2 - \min_{\substack{p=1,2,\dots,K \\ j=1,2,\dots,p \\ j \neq i}} \|x_a^{(i)} - x_p^{(i)}\|_2]_+ \quad (1)$$

式中: $x_a^{(i)}$ 、 $x_p^{(i)}$ 、 $x_n^{(i)}$ 分别为从锚样本、正样本以及负样本提取的特征表示; P 代表每训练批次随机抽取的行人身份数量; K 表示每个身份包含的实例数量。该部分模型的整体损失函数可表示为

$$L_{\text{ssg}} = L_r(f_r, y_r) + L_{\text{triple}}(f_{t_up}, y_{t_up}) + L_{\text{triple}}(f_{t_low}, y_{t_low}) + L_{\text{triple}}(f_{ie}, y_{ie}) \quad (2)$$

式中: f_{ie} 为整体行人特征输入全连接层后得到的整体嵌入向量, 与整体行人特征具有相同的标签信息。

1.2 基于多尺度特征转换的风格转换模型

SPGAN 网络是一个用于无监督行人重识别的风格转换网络, 它以 CycleGAN 为基础框架, 首先, 通过 2 个镜像对称的生成对抗网络 (generative adversarial network, GAN) 形成一个环形网络, 来进行图像的风格转换; 其次, 加入了新的损失函数, 该损失函数以目标域的身份作为约束, 辅助生成器去生成目标域风格的行人图片; 另外, 该模型为了保证行人图片转换过程中不丢失图片身份相关的信息, 加入了孪生网络 SiaNet 以及对应的一致性损失函数, 保证行人图片经过转换后其相关身份信息不丢失。本文提出的基于多尺度特征转换的风格转换模型由多尺度特征转换的生成器、判别器以及 SiaNet 网络组成。

1.2.1 多尺度特征转换的生成器模型

SPGAN 的生成器是编码器-转换器-解码器体系结构, 其编码器的作用是提取输入图像的特征信息; 其转换器由原始残差块组成, 作用是将源域输入图像特征转换为接近于目标域图像风格的特征; 其解码器的作用是将特征向量进行复原。但是该生成器存在问题, 首先, 由于编码器下采样和解码器的上采样过程中, 通过的特征具有尺度和位置不变性, 而且在 SPGAN 的生成器中, 只在一个尺度上使用原始残差块, 只有最低层的特征信息可以进行特征的转换, 这样尺度单一的特征信息转换会导致部分信息的丢失以及网络学习功能的下降; 其次, 由于原始残差块的结构原因, 特征信息在通过最低层的原始残差块时, 原始残差块限制可以通过瓶颈的信息, 这样会限制网络学习的功能。针对上述问题, 本文对 SPGAN 的生成器进行改进。

改进的生成器模型如图 3 所示, ReLU 为线性

整流函数, Tanh 为双曲正切函数。



图 3 改进生成器模型

Fig. 3 Improved generator model

首先, 采用一种改进残差块模块^[13], 替换原来承担转换作用的原始残差块。原始残差块如图 4(a) 所示, 其结构顺序一般为: 使用 Conv + BN + ReLU + Conv + BN 模式作为侧边分支的模式, 特征通过侧边分支与原来的特征相加, 再通过 ReLU 层传递信息。改进残差块如图 4(b), 该模块与原始残差块相比, 改变了原有的结构顺序, 改进残差块采用 BN + ReLU + Conv + BN + ReLU + Conv 模式作为侧边分支, 特征通过侧边分支直接与原特征相加。将主干的 ReLU 层直接放到侧边分支中, 并且改变其排列顺序, 将 ReLU 和 BN 层作为权重层的预激活层, 这种结构有利于更多的特征信息通过, 保证特征信息在传递过程中更好地防止特征信息丢失, 且生成器更易于训练; 其次, 针对生成器结构中特征转换尺度单一问题, 本文将改进残差块引入到生成器上采样及下采样的对应层中, 在此结构中, 生成器在上采样以及下采样过程中均有改进残差块承担特征信息转换作用, 在防止特征信息在传递过程中丢失的前提下, 形成一个多尺度特征转换的生成器, 一定程度上解决了原始 SPGAN 在下采样和上采样过程中特征信息无法进行多尺度转换以及特征信息丢失的问题, 提升生成器网络的学习性能, 进一步提升改进模型在风格转换上的效果。

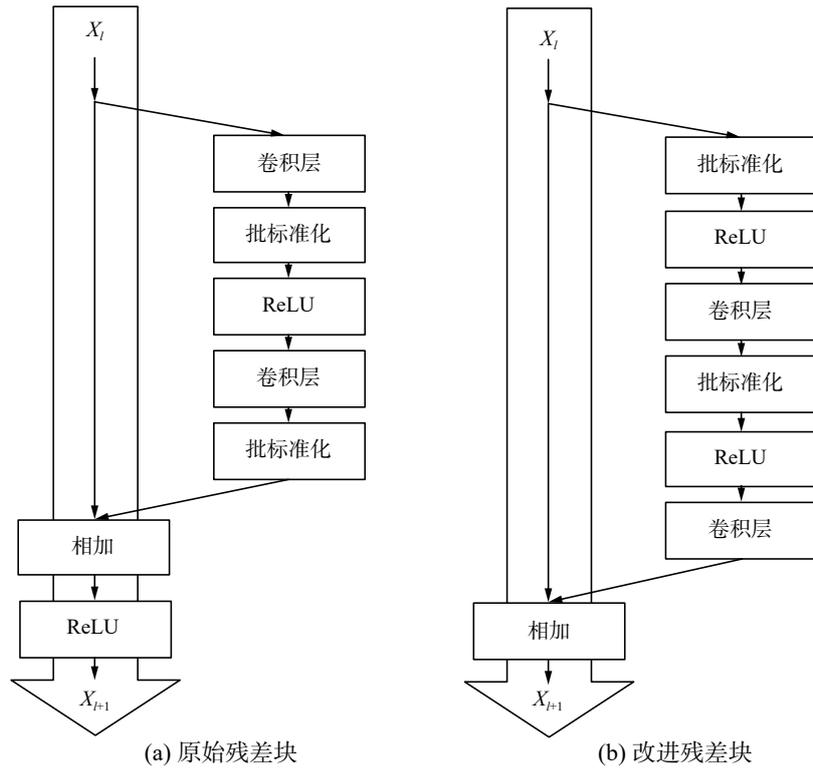


图 4 生成器残差块示意
Fig. 4 Schematic diagram of generator residuals

1.2.2 判别器模型

本文的判别器属于卷积网络,判别器将图像作为输入对其进行特征提取,通过判别提取特征的真假来预测输入图像是原始图像还是生成图像。

1.2.3 SiaNet 网络

SiaNet 网络通过利用 Contrastive Loss^[14] 进行监督训练,通过训练, SiaNet 网络将风格转换后的图片与其转换前图片拉近,将风格迁移后的图片与目标数据集中的其他图片推远,这样可以保证行人的 ID 信息不变。

1.2.4 损失函数

该部分模型包含 2 个生成器—判别器对,分别是 {G, D_T} 和 {F, D_S}, 它们将源域(目标域)的图像转换成目标域(源域)风格的图像。对于生成器和对应的目标域判别器,其对抗损失函数可表示为

$$L_{Tadv}(G, D_T, p_x, p_y) = E_{y \sim p_y} [(D_T(y) - 1)^2] + E_{x \sim p_x} [(D_T(G(x)))^2] \quad (3)$$

式中 p_x 和 p_y 分别表示源域和目标域的样本分布。

对于生成器 F 和对应的源域判别器 D_S , 其对抗损失函数可表示为

$$L_{Sadv}(F, D_S, p_y, p_x) = E_{x \sim p_x} [(D_S(x) - 1)^2] + E_{y \sim p_y} [(D_S(F(y)))^2] \quad (4)$$

由于没有成对数据进行训练,本模型通过使

用正向转换以及反向转换的循环后恢复原始图像,以减少可能的映射函数空间,其循环一致性损失函数可表示为

$$L_{cyc}(G, F) = E_{x \sim p_x} [\|F(G(x) - x)\|_1] + E_{y \sim p_y} [\|G(F(y) - y)\|_1] \quad (5)$$

除了使用对抗损失函数以及循环一致性损失函数,为了保证生成器生成图像的稳定性,本模型还使用目标域身份约束作为图像——图像转换的辅助方法,该身份约束损失函数可表示为

$$L_{ide}(G, F, p_x, p_y) = E_{x \sim p_x} \|F(x) - x\|_1 + E_{y \sim p_y} \|G(y) - y\|_1 \quad (6)$$

该部分模型的整体损失函数可表示为

$$L_{sp} = L_{Tadv} + L_{Sadv} + \lambda_1 L_{cyc} + \lambda_2 L_{ide} + \lambda_3 L_{con} \quad (7)$$

其中, $\lambda_1=10, \lambda_2=5, \lambda_3=2, m=10$ 。

2 仿真实验与结果分析

2.1 实验环境和参数设置

本实验的硬件环境配置为 Intel©Core™i7 9700K 64 位处理器,采用 2 个 GeForce GTX 2080 Ti GPU 进行运算加速,操作系统为 64 位 Ubuntu 16.04,采用基于 python3.6 以及 Pytorch 1.0.0 完成程序编程。

基于多尺度特征转换的风格转换模型的迭代次数为 15,批量大小为 8;基于风格转换的无监督聚类行人重识别模型的迭代次数为 30,每次迭代

包括 70 个 epoch, 批量大小为 128。

2.2 数据集设置

本文采用目前无监督行人重识别领域常用的公开数据集 Market1501^[15] 以及 Duke-MTMC-reID^[16] 进行实验。文献 [17-19] 模型均需要使用源域数据进行模型预训练, 然后使用目标域数据进行正式训练及测试, 本文数据集的设置与目前无监督行人重识别采用的数据集设置标准一致。

2.3 仿真实验与结果分析

实验从有效性实验以及与主流模型的对比实验 2 个方面对模型效果进行验证。在实验测试中, 将测试集分为查询集与候选集, 针对每个查询行人图像计算其余所有候选图像的距离并按照距离大小从小到大的顺序对候选图像排序。实验

采用无监督行人重识别通用的累积匹配特性 (cumulative match characteristic curve, CMC) 中的 Rank-1、Rank-5、Rank-10 以及平均精度 (mAP) 作为评价指标。

2.3.1 算法有效性验证实验

为了验证所提出的模型的有效性, 实验在与目前无监督行人重识别工作一致的 2 种通用跨域设定上进行实验。它们分别是以 Market1501 为源域, 在 Duke-MTMC-reID 上进行无监督实验; 以 Duke-MTMC-reID 为源域, 在 Market1501 进行无监督实验。

首先, 对 2.1 节提出的改进算法——基于风格转换的 SSG 模型进行实验。实验对该模型进行有效性实验验证。实验结果如表 1 所示。

表 1 改进 SSG 有效性实验验证结果

Table 1 Results of improved SSG verification experiment

%

| 算法 | Market1501→Duke-MTMC-reID | | | | Duke-MTMC-reID→Market1501 | | | |
|-----|---------------------------|-------------|-------------|-------------|---------------------------|-------------|-------------|-------------|
| | Rank-1 | Rank-5 | Rank-10 | mAP | Rank-1 | Rank-5 | Rank-10 | mAP |
| SSG | 67.8 | 78.8 | 82.0 | 48.7 | 74.2 | 86.5 | 89.9 | 51.9 |
| 本文 | 68.3 | 78.6 | 82.7 | 49.0 | 78.4 | 89.8 | 92.7 | 57.3 |

表 1 中的数据为模型在目标域上进行测试的结果, 表中的模型均使用源域数据进行预训练, 使用目标域数据进行正式训练及测试。从表中可以看出, 将 SPGAN 模型引入到 SSG 模型中后, 其识别效果优于 SSG 模型, 在以 Duke-MTMC-reID 数据集为目标域的测试实验中, 仅有 Rank-5 指标低于 SSG 模型 0.2%, 其他指标均有不同幅度提升。实验结果证明该部分实验的有效性。

同时, 实验结果说明无监督聚类的行人重识

别模型性能受到源域和目标域差异性过大带来的影响, SSG 模型通过聚类方法挖掘目标域数据间的相似性, 只能降低部分域间差异性对模型识别效果的影响。因此, 本部分实验证明引入风格转换进一步降低域间差异性带来的影响是有必要的。

接下来, 对改进的风格转换模型——基于多尺度特征转换的行人重识别风格转换模型—SPGAN+进行有效性实验。SPGAN+图片转换效果及对比效果如图 5、6 所示, 实验结果如表 2 所示。



图 5 原数据集抽样图片和 SPGAN+转换的抽样图片

Fig. 5 Sample images of original data and transferred data using the SPGAN+



图 6 图像转换的抽样图像

Fig. 6 Visual examples of image-image translation.

表 2 改进 SPGAN 有效性实验验证结果

Table 2 Results of improved SPGAN verification experiment

%

| 算法 | Market1501→Duke-MTMC-reID | | | | Duke-MTMC-reID→Market1501 | | | |
|--------|---------------------------|-------------|-------------|-------------|---------------------------|-------------|-------------|-------------|
| | Rank-1 | Rank-5 | Rank-10 | mAP | Rank-1 | Rank-5 | Rank-10 | mAP |
| 基线模型 | 33.7 | 49.2 | 55.3 | 18.8 | 48.1 | 64.0 | 70.8 | 21.3 |
| SPGAN | 39.5 | 56.1 | 62.5 | 23.4 | 53.6 | 71.1 | 77.1 | 25.0 |
| SPGAN+ | 43.1 | 57.9 | 64.0 | 25.1 | 54.8 | 71.8 | 78.2 | 26.3 |

图 5 为原数据集抽样图片和 SPGAN+转换的抽样图片。由图 5 可以看出, 本文模型在两域间的风格转换上达到较好效果, 经过 SPGAN+模型转换后的源域图片风格明显接近于目标域图片风格, 实验证明, 改进的风格转换模型在视觉效果上的有效性。

图 6 中, 图 6(a) 表示原数据集的随机选取图片, 图 6(b) 表示使用 SPGAN 模型将源域图像转换为目标域图像的随机选取图片, 图 6(c) 表示使用本文改进的 SPGAN+模型将源域图像转换为目标域图像的随机选取图片。经过对比, 可以看出使用 SPGAN+模型进行风格转换后的图像相比 SPGAN 模型更接近于目标域风格。

从视觉效果上来说, SPGAN+模型的风格转换效果优于 SPGAN 模型, 也说明通过本文方法改进了生成器中特征转换尺度单一及特征信息传递效率低下问题, 明显提升该模型的生成器的学习性能, 提升模型风格转换的视觉效果。

表 2 中, Baseline 表示直接使用源域数据源训练的 Baseline 模型在目标域上进行测试的实验结果, SPGAN 表示首先将源域图像通过 SPGAN 模型转换成目标域风格的图像, 然后训练 Baseline

模型, 在目标域上进行测试实验; SPGAN+代表本文的改进行人重识别风格转换模型, 首先将源域图像通过 SPGAN+模型转换成目标域风格的图像, 然后训练 Baseline 模型, 并在目标域上进行测试的实验结果。由表 2 可知, 本实验改进的 SPGAN+模型的性能有较好提升, 在以 Market1501 以及 Duke-MTMC-reID 数据集为目标域的测试中, 改进模型的各项实验指标均有较好的提升, 验证了改进的 SPGAN+模型的有效性。

实验结果说明, SPGAN 模型的生成器中, 其生成器结构存在特征转换尺度单一的缺点, 且其承担转换作用的原始残差块存在信息传递效率低和限制可以通过瓶颈的特征信息的缺点, 导致生成器的学习性能不佳, 使 SPGAN 模型在风格转换上不能达到较好效果。因此, 本部分实验证明本文改进算法 SPGAN+的生成器通过解决转换尺度单一以及特征传递效率低下的问题可以让该模型得到更好的域间风格转换效果。

最后, 进行基于风格转换的无监督聚类行人重识别算法的实验。实验对整体模型与 SSG 以及第 1 部分实验进行有效性对比验证。实验结果如表 3 所示。

表 3 改进 SSG 有效性实验验证结果
Table 3 Results of improved SSG verification experiment

%

| 算法 | Market1501→Duke-MTMC-reID | | | | Duke-MTMC-reID→Market1501 | | | |
|------------|---------------------------|-------------|-------------|-------------|---------------------------|-------------|-------------|-------------|
| | Rank-1 | Rank-5 | Rank-10 | mAP | Rank-1 | Rank-5 | Rank-10 | mAP |
| SSG | 67.8 | 78.8 | 82.0 | 48.7 | 74.2 | 86.5 | 89.9 | 51.9 |
| SSG+SPGAN | 68.3 | 78.6 | 82.7 | 49.0 | 78.4 | 89.8 | 92.7 | 57.3 |
| SSG+SPGAN+ | 68.0 | 79.6 | 83.3 | 49.1 | 79.4 | 89.9 | 93.2 | 57.4 |

表 3 中, SSG+SPGAN 表示本实验通过将 SPGAN 引入到 SSG 模型中进行实验的结果; SSG+SPGAN+表示本实验通过将 SPGAN+引入到 SSG 模型中进行实验的结果。由表 3 可知, 将风格转换模型 SPGAN 以及 SPGAN+分别引入到 SSG 模型中, 最终识别效果均好于只使用 SSG 模型的识别效果; SSG+SPGAN 与 SSG+SPGAN+实验结果, 后者在性能上优于前者, 后者的识别率仅在以 Duke-MTMC-reID 为目标域的实验中, 其 rank-1 略低于前者, 其他项识别效率均高于前者。

这部分实验说明, 针对 SPGAN 模型生成器转换尺度单一以及特征信息传递效率低的问题的改

进, 在整体模型中对模型最终识别率存在提升效果, 进一步验证了改进风格转换模型 SPGAN+的有效性以及整体模型的有效性。

2.3.2 本文算法与主流算法对比实验

实验对目前各个方向最先进的基于上下文呈现的行人重识别网络 (context rendering network, CR-GAN)^[17]、基于存储器的领域自适应的行人重识别 (exemplar memory for domain adaptive person re-identification, ECN)^[18]、于深度软多标签参考学习的行人重识别 (deep soft multilabel reference learning, MAR)^[19] 和 SSG 进行对比, 实验结果如表 4 所示。

表 4 本文算法与先进无监督行人重识别算法结果对比

Table 4 Comparison of the results between our algorithm and the advanced unsupervised person re-identification algorithms

%

| 算法 | Market1501→Duke-MTMC-reID | | | | Duke-MTMC-reID→Market1501 | | | |
|--------|---------------------------|-------------|-------------|-------------|---------------------------|-------------|-------------|-------------|
| | Rank-1 | Rank-5 | Rank-10 | mAP | Rank-1 | Rank-5 | Rank-10 | mAP |
| CR-GAN | 52.2 | — | — | 30.0 | 59.6 | — | — | 29.6 |
| MAR | 67.1 | 79.8 | — | 48.0 | 67.7 | 81.9 | — | 48.0 |
| ECN | 63.3 | 75.8 | 80.4 | 40.4 | 75.1 | 87.6 | 91.6 | 43.0 |
| SSG | 67.8 | 78.8 | 82.0 | 48.7 | 74.2 | 86.5 | 89.9 | 51.9 |
| 本文算法 | 68.0 | 79.6 | 83.3 | 49.1 | 79.4 | 89.9 | 93.2 | 57.4 |

在表 4 中, “—”表示对应方法未在数据集上实现。由表 4 可知, 本实验提出的方法相较于其他主流方法, 本实验的方法仅在 Duke-MTMC-reID 数据集的 rank-5 上略低于 MAR, 其他指标均取得了最高精度, 说明本文算法优于上述对比算法。

3 结束语

首先, 本文发现基于聚类的无监督行人重识别算法—SSG 中源域和目标域之间数据分布的差异性是影响模型性能的重要因素, 而减小域间风格差异性的方法可以降低该问题对 SSG 模型识别效果的影响。其次, 本文发现引入的 SPGAN 模型的生成器存在特征转换尺度单一以及特征信息传递效率较低的问题, 该问题影响 SPGAN 模型的风格转换效果, 降低整体模型的识别

率, 通过实现特征的多分辨率尺度转换, 并且提升生成器特征传递效率, 可以提升 SPGAN 模型的风格转换效果, 进一步提升最终识别效果。在目前无监督行人重识别领域使用的通用数据集设定上, 对本实验提出的算法进行验证, 本文实验结果表明, 将 SPGAN 引入到 SSG 模型中后, 模型识别效果得到提升; 对引入的 SPGAN 模型生成器进行改进, 提升其转换性能, 模型的识别效果得到进一步提升。

无监督聚类行人重识别模型 SSG 中使用的聚类方法还存在一些问题, 提升 SSG 模型中对特征聚类的准确度是今后需要完善的工作。

参考文献:

- [1] 刘帅师, 程曦, 郭文燕, 等. 深度学习研究方法研究新进展 [J].

- 智能系统学报, 2016, 11(5): 567–577.
- LIU Shuaishi, CHENG Xi, GUO Wenyan, et al. Progress report on new research in deep learning[J]. *CAAI transactions on intelligent systems*, 2016, 11(5): 567–577.
- [2] 宋婉茹, 赵晴晴, 陈昌红, 等. 行人重识别研究综述 [J]. *智能系统学报*, 2017, 12(6): 770–780.
- SONG Wanru, ZHAO Qingqing, CHEN Changhong, et al. Survey on pedestrian re-identification research[J]. *CAAI transactions on intelligent systems*, 2017, 12(6): 770–780.
- [3] GENG Mengyue, WANG Yaowei, XIANG Tao, et al. Deep transfer learning for person re-identification[J]. arXiv preprint arXiv: 1611.05244, 2016: 00007480.
- [4] LI Wei, ZHAO Rui, XIAO Tong, et al. Deepreid: deep filter pairing neural network for person re-identification[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA, 2014: 152–159.
- [5] ZHONG Zhun, ZHENG Liang, LI Shaozi, et al. Generalizing a person retrieval model hetero- and homogeneously[C]// Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany, 2018: 176–192.
- [6] PENG Peixi, XIANG Tao, WANG Yaowei, et al. Unsupervised cross-dataset transfer learning for person re-identification[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA, 2016: 1306–1315.
- [7] DENG Weijian, ZHENG Liang, YE Qixiang, et al. Image-image domain adaptation with preserved self-similarity and domain-dissimilarity for person re-identification[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA, 2018: 994–1003.
- [8] FU Yang, WEI Yunchao, WANG Guanshuo, et al. Self-similarity grouping: a simple unsupervised cross domain adaptation approach for person re-identification[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South), 2019: 6111–6120.
- [9] ZHU Junyan, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]// Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy, 2017: 2242–2251.
- [10] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA, 2016: 770–778.
- [11] ESTER M, KRIEGEL H P, SANDER J, et al. A density-based algorithm for discovering clusters in large spatial databases with noise[C]// Proceedings of the Second International Conference on Knowledge Discovery and Data Mining. Portland, Oregon, USA, 1996: 226–231.
- [12] HERMANS A, BEYER L, LEIBE B. In defense of the triplet loss for person re-identification[J]. arXiv preprint arXiv: 1703.07737, 2017.
- [13] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Identity mappings in deep residual networks[C]//14th European Conference on Computer Vision. Amsterdam, The Netherlands, 2016: 630–645.
- [14] HADSELL R, CHOPRA S, LECUN Y. Dimensionality reduction by learning an invariant mapping[C]//2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York, NY, USA, 2006: 1735–1742.
- [15] ZHENG Liang, SHEN Liyue, TIAN Lu, et al. Scalable person re-identification: a benchmark[C]// Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile, 2015: 1116–1124.
- [16] RISTANI E, SOLERA F, ZOU R, et al. Performance measures and a data set for multi-target, multi-camera tracking[C]// European Conference on Computer Vision. Amsterdam, The Netherlands, 2016: 17–35.
- [17] CHEN Yanbei, ZHU Xiatian, GONG Shaogang. Instance-guided context rendering for cross-domain person re-identification[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South), 2019: 232–242.
- [18] ZHONG Zhun, ZHENG Liang, LUO Zhiming, et al. Invariance matters: exemplar memory for domain adaptive person re-identification[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA, 2019: 598–607.
- [19] YU Hongxing, ZHENG Weishi, WU Ancong, et al. Unsupervised person re-identification by soft multilabel learning[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA, 2019: 2143–2152.

作者简介:



张智, 硕士研究生, 主要研究方向为深度学习、图像处理。



毕晓君, 教授, 博士生导师, 主要研究方向为信息智能处理、数字图像处理、智能优化算法及机器学习。主持国家自然科学基金面上项目 2 项、科技部国际合作项目面上项目 1 项、教育部博士点基金项目 1 项、工业和信息化部海洋工程装备科研项目子项目 1 项、民品横向课题 1 项, 获国家专利 8 项。发表学术论文 170 余篇, 出版学术专著 3 部。